

ARCH DEVELOPMENT CORPORATION

October 11, 1995  
A21911PC DO/AD/al/cp

5

CLAIMS

1. A method for determining the sequence of a nucleic acid molecule,  
comprising the steps of:
  - (a) identifying sequences from the molecule by:
    - (i) hybridizing the molecule to complementary sequences of oligonucleotides from two sets of small oligonucleotide probes of known sequence, wherein the first set of probes are attached to a solid support and the second set of probes are labelled probes in solution; and
    - (ii) covalently bonding a hybridized oligonucleotide from said first set of probes to a hybridized oligonucleotide from said second set of probes;
  - (b) identifying overlapping stretches of sequence from the sequences identified in step (a); and
  - (c) assembling the nucleic acid sequence of the molecule from said overlapping sequences identified.
2. The method of claim 1, wherein said hybridization is carried out in cycles.

3. A method for determining the sequence of a nucleic acid molecule, comprising the steps of:
- (a) fragmenting the nucleic acid molecule to be sequenced to provide intermediate length nucleic acid fragments;
- (b) identifying sequences from said fragments by:
- (i) hybridizing the fragments to complementary sequences of oligonucleotides from two sets of small oligonucleotide probes of known sequence, wherein the first set of probes are attached to a solid support and the second set of probes are labelled probes in solution; and
- (ii) covalently bonding a hybridized oligonucleotide from said first set of probes to a hybridized oligonucleotide from said second set of probes;
- (c) identifying overlapping stretches of sequence from said sequences identified in step (b); and
- (d) assembling the nucleic acid sequence of the molecule from said overlapping sequences identified.
4. The method of claim 3, wherein said fragments are sequentially hybridized to complementary sequences from two sets of small oligonucleotide probes of known sequence.

5. The method of claim 3, wherein said fragments are simultaneously hybridized to complementary sequences from two sets of small oligonucleotide probes of known sequence.
- 5 6. The method of claim 3, wherein said intermediate length nucleic acid fragments are between about 10 nucleotides and about 40 nucleotides in length and said small oligonucleotides probes are between about 4 nucleotides and about 9 nucleotides in length.
- 10 7. The method of claim 3, wherein said oligonucleotide probes hybridize to completely complementary sequences from said fragments.
8. The method of claim 3, wherein said oligonucleotide probes hybridize to immediately adjacent sequences from said fragments.
- 15 9. The method of claim 8, wherein said oligonucleotide probes hybridize to completely complementary and immediately adjacent sequences from said fragments.
- 20 10. The method of claim 3, wherein said oligonucleotide probes are covalently bonded by enzymatic ligation.
11. The method of claim 3, wherein said oligonucleotide probes are covalently bonded using a chemical ligating agent.
- 25 12. The method of claim 3, wherein step (b) comprises the steps of:
  - (a) contacting said first set of small attached oligonucleotide probes with said intermediate length nucleic acid fragments under hybridization conditions effective to allow only those fragments

with a completely complementary sequence to hybridize to a probe, thereby forming primary complexes wherein the fragment has hybridized and free sequences;

- 5                     (b) contacting said primary complexes with said second set of small labelled oligonucleotide probes under hybridization conditions effective to allow only those probes with completely complementary sequences to hybridize to a free fragment sequence, thereby forming secondary complexes wherein the fragment is hybridized to an attached probe and a labelled probe;
- 10                   (c) covalently bonding said attached probe and said labelled probe;
- 15                   (d) removing from said secondary complexes labelled probes that are not covalently bonded to an attached probe, thereby forming covalently bonded complexes;
- 20                   (e) detecting said covalently bonded complexes by detecting the presence of the label; and
- 25                   (f) identifying sequences from the nucleic acid fragments in said covalently bonded complexes by connecting the known sequences of the hybridized attached and labelled probes.

25       13. A method of nucleic acid sequencing comprising the steps of:

- (a) fragmenting the nucleic acid to be sequenced to provide nucleic acid fragments of length T;

- 5

(b) preparing an array of immobilized oligonucleotide probes of known sequences and length F and a set of labelled oligonucleotide probes in solution of known sequences and length P, wherein  $F + P \leq T$ ;

10

(c) contacting said array of immobilized oligonucleotide probes with said nucleic acid fragments under hybridization conditions effective to allow the formation of primary complexes with hybridized, completely complementary sequences of length F and non-hybridized fragment sequences of length  $T - F$ ;

15

(d) contacting said complexes with said set of labelled oligonucleotide probes under hybridization conditions effective to allow only the formation of secondary complexes with hybridized, completely complementary sequences of length F and immediately adjacent hybridized, completely complementary sequences of length P;

20

(e) covalently bonding said labelled oligonucleotide probes to said immediately adjacent immobilized oligonucleotide probes;

25

(f) detecting said secondary complexes by detecting the presence of the label;

(g) identifying sequences of length  $F + P$  from the nucleic acid fragments in said secondary complexes by combining the known sequences of the hybridized immobilized and labelled probes;

30

(h) determining stretches of said sequences of length  $F + P$  that overlap; and

- (i) assembling the complete nucleic acid sequence from said overlapping sequences.
14. The method of claim 13, wherein length T is about three times  
5 longer than length F.
15. The method of claim 13, wherein length T is between about 10  
nucleotides and about 40 nucleotides, length F is between about 4  
10 nucleotides and about 9 nucleotides and length P is between about  
4 nucleotides and about 9 nucleotides.
16. The method of claim 15, wherein length T is about 20 nucleotides,  
length F is about 6 nucleotides and length P is about 6 nucleotides.
- 15 17. The method of claim 13, wherein said immediately adjacent immobilized and labelled oligonucleotide probes are covalently bonded by enzymatic ligation.
- 20 18. The method of claim 13, wherein said immediately adjacent immobilized and labelled oligonucleotide probes are covalently bonded using a chemical ligating agent.
19. A method of nucleic acid sequencing comprising the steps of:
- 25 (a) fragmenting the nucleic acid to be sequenced to provide intermediate length nucleic acid fragments;
- (b) contacting an array of immobilized small oligonucleotide probes of known sequences with said nucleic acid fragments under hybridization conditions effective to allow only those fragments
- 30

with a completely complementary sequence to hybridize to a probe, thereby forming primary complexes wherein the fragment has hybridized and non-hybridized sequences;

- 5                     (c) contacting said primary complexes with a set of labelled small oligonucleotide probes in solution of known sequences under hybridization conditions effective to allow only those probes with completely complementary sequences to hybridize to a non-hybridized fragment sequence, thereby forming secondary complexes wherein the fragment is hybridized to an immobilized probe and a labelled probe;
- 10                   (d) covalently bonding said labelled oligonucleotide probes to said immediately adjacent immobilized oligonucleotide probes;
- 15                   (e) removing from said secondary complexes labelled probes that are not covalently bonded to an immobilized probe, thereby forming covalently bonded complexes;
- 20                   (f) detecting said covalently bonded complexes by detecting the presence of the label;
- 25                   (g) identifying sequences from the nucleic acid fragments in said covalently bonded complexes by combining the known sequences of the hybridized immobilized and labelled probes;
- 30                   (h) determining stretches of said sequences that overlap; and
- (i) assembling the complete nucleic acid sequence from said overlapping sequences identified.

20. The method of claim 19, wherein the nucleic acid is cloned DNA or chromosomal DNA.
21. The method of claim 19, wherein the nucleic acid is mRNA.
- 5 22. The method of claim 19, wherein the nucleic acid is fragmented by restriction enzyme digestion, ultrasound treatment, NaOH treatment or low pressure shearing.
- 10 23. The method of claim 19, wherein the nucleic acid fragments are between about 10 nucleotides and about 100 nucleotides in length.
24. The method of claim 19, wherein the oligonucleotide probes are between about 4 nucleotides and about 9 nucleotides in length.
- 15 25. The method of claim 24, wherein the oligonucleotide probes are about 6 nucleotides in length.
26. The method of claim 19, wherein said immobilized oligonucleotides are attached to a glass, polystyrene or teflon solid support.
- 20 27. The method of claim 19, wherein said immobilized oligonucleotides are attached to a solid support via a phosphodiester linkage.
- 25 28. The method of claim 19, wherein said immobilized oligonucleotides are attached to a solid support via a light-activated synthetic mechanism.
29. The method of claim 19, wherein the labelled oligonucleotide probes are labelled with a non-radioactive isotope or a fluorescent dye.

30. The method of claim 19, wherein the labelled oligonucleotide probes are labelled with  $^{35}\text{S}$ ,  $^{32}\text{P}$  or  $^{33}\text{P}$ .
31. The method of claim 19, wherein said nucleic acid fragment or one of said oligonucleotide probes contains a modified base or a universal base.
32. The method of claim 19, wherein labelled probes that are not covalently bonded to an immobilized probe are removed from the secondary complexes by stringent washing conditions.
33. The method of claim 19, wherein said immediately adjacent probes are chemically bonded.
34. The method of claim 19, wherein said immediately adjacent probes are ligated enzymatically.
35. The method of claim 19, wherein multiple arrays of immobilized oligonucleotides are arranged in the form of a sequencing chip.
36. A method of nucleic acid sequencing comprising the steps of:
  - (a) fragmenting the nucleic acid to be sequenced to provide nucleic acid fragments of between about 10 nucleotides and about 40 nucleotides in length;
  - (b) contacting an array of immobilized oligonucleotide probes with known sequences of between about 4 nucleotides and about 9 nucleotides in length with said nucleic acid fragments under hybridization conditions effective to allow only those fragments

with a completely complementary sequence to hybridize to a probe, thereby forming primary complexes wherein the fragment has hybridized and non-hybridized sequences;

- 5       (c) contacting said complexes with a set of  $^{32}\text{P}$ -labelled or  $^{33}\text{P}$ -labelled oligonucleotide probes with known sequences of between about 4 nucleotides and about 9 nucleotides in length under hybridization conditions effective to allow only those labelled probes with completely complementary sequences to hybridize to a non-hybridized fragment sequence, thereby forming secondary complexes wherein the fragment is hybridized to an immobilized probe and a  $^{32}\text{P}$ -labelled or  $^{33}\text{P}$ -labelled probe;
- 10      (d) ligating the immobilized probes and labelled probes that are immediately adjacent with a DNA ligase enzyme, thereby forming ligated secondary complexes;
- 15      (e) removing from the secondary complexes any non-ligated labelled probes;
- 20      (f) detecting said ligated secondary complexes by detecting the presence of the  $^{32}\text{P}$  or  $^{33}\text{P}$  label;
- 25      (g) identifying sequences from the nucleic acid fragments in said ligated secondary complexes by combining the known sequences of the ligated probes;
- (h) determining stretches of said sequences that overlap; and

- (i) assembling the complete nucleic acid sequence from said overlapping sequences.
37. A kit for use in nucleic acid sequencing, comprising a solid support chip having attached an arrangement of oligonucleotide probes of known sequences, said oligonucleotides being capable of taking part in hybridization reactions, a set of containers comprising solutions of labelled oligonucleotide probes of known sequences, and a ligating agent.
- 10
38. The kit of claim 37, wherein multiple chips of immobilized oligonucleotide probes are arranged in the form of a sequencing array.
- 15
39. The kit of claim 37, wherein the oligonucleotide probes are between about 4 nucleotides and about 9 nucleotides in length.
- 20
40. The kit of claim 39, wherein the oligonucleotide probes are about 6 nucleotides in length.
- 25
41. The kit of claim 37, wherein the oligonucleotide probes are attached to a glass, polystyrene or teflon solid support.
42. The kit of claim 37, wherein the oligonucleotide probes are attached to a solid support via a phosphodiester linkage.
- 25
43. The kit of claim 37, wherein the oligonucleotide probes are attached to a solid support via a light-activated synthetic mechanism.
- 30
44. The kit of claim 37, wherein the labelled oligonucleotide probes are labelled with a non-radioactive isotope or a fluorescent dye.

45. The kit of claim 37, wherein one of the oligonucleotide probes contains a modified or a universal base.
46. The kit of claim 37, wherein the labelled oligonucleotide probes are labelled with  $^{35}\text{S}$ ,  $^{32}\text{P}$  or  $^{33}\text{P}$ .
47. The kit of claim 37, wherein said ligating agent is a chemical ligating agent.
- 10 48. The kit of claim 37, wherein the ligating agent is a DNA ligase enzyme.

add  
a3